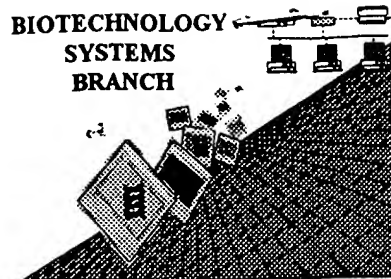


## **RAW SEQUENCE LISTING** **ERROR REPORT**



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/807,223  
Source: PA/09  
Date Processed by STIC: 4/24/2001

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

**FOR CRE SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.**

**FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.**

**PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW**

### **Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§ 1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST 25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).

Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

**Checker Version 3.0 can be down loaded from the USPTO website at the following address:**

**<http://www.uspto.gov/web/offices/pac/checker>**

BEST AVAILABLE COPY

# Raw Sequence Listing Error Summary

## ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER: 09/807,223

**ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE**

- 1        Wrapped Nucleics      The number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 2        Wrapped Aminos      The amino acid number/text at the end of each line "wrapped " down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 3        Incorrect Line Length      The rules require that a line not exceed 72 characters in length. This includes spaces.
- 4        Misaligned Amino Acid Numbering      The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.
- 5        Non-ASCII      This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.  
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
- 6        Variable Length      Sequence(s)        contain n's or Xaa's which represented more than one residue.  
As per the rules, each n or Xaa can only represent a single residue.  
Please present the maximum number of each residue having variable length and indicate in the (ix) feature section that some may be missing.
- 7        PatentIn ver. 2.0 "bug"      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s)       . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. **This applies primarily to the mandatory <220>-<223> sections for Artificial or Unknown sequences.**
- 8        Skipped Sequences (OLD RULES)      Sequence(s)        missing. If intentional, please use the following format for each skipped sequence:  
**(2) INFORMATION FOR SEQ ID NO:X:**  
**(i) SEQUENCE CHARACTERISTICS:**(Do not insert any headings under "SEQUENCE CHARACTERISTICS")  
**(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:**  
**This sequence is intentionally skipped**  
  
Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
- 9        Skipped Sequences (NEW RULES)      Sequence(s)        missing. If intentional, please use the following format for each skipped sequence.  
**<210> sequence id number**  
**<400> sequence id number**  
**000**
- 10        Use of n's or Xaa's (NEW RULES)      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 11        Use of "Artificial" (NEW RULES)      Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules.  
Valid response is Artificial Sequence.
- 12        Use of <220>Feature (NEW RULES)      Sequence(s)        are missing the <220>Feature and associated headings.  
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial Sequence" or "Unknown"  
**Please explain source of genetic material in <220> to <223> section.**  
**(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)**
- 13        PatentIn ver. 2.0 "bug"      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).  
Instead, please use "File Manager" or any other means to copy file to floppy disk.

PCT

## RAW SEQUENCE LISTING

DATE: 04/24/2001

PATENT APPLICATION: US/09/807,223

TIME: 13:26:16

Input Set : A:\Q63988 Sequence Listing.txt

Output Set: N:\CRF3\04242001\I807223.raw

Does Not Comply  
Corrected Diskette Needed

3 <110> APPLICANT: Sumitomo Pharmaceutical Co., Ltd.  
 5 <120> TITLE OF INVENTION: RECOMBINASE-EXPRESSING CELLS  
 7 <130> FILE REFERENCE: Q63988  
 OK 9 <140> CURRENT APPLICATION NUMBER: US/09/807,223  
 9 <141> CURRENT FILING DATE: 2001-04-11  
 9 <150> PRIOR APPLICATION NUMBER: JP 10-289785  
 10 <151> PRIOR FILING DATE: 1998-10-12  
 12 <160> NUMBER OF SEQ ID NOS: 5  
 14 <170> SOFTWARE: PatentIn version 3.0  
 16 <210> SEQ ID NO: 1  
 17 <211> LENGTH: 54  
 18 <212> TYPE: DNA  
 C--> 19 <213> ORGANISM: Artificial *see item 11 on Error Summary Sheet*  
 21 <220> FEATURE:  
 22 <223> OTHER INFORMATION: FLP recognition sequence  
 24 <400> SEQUENCE: 1  
 25 aaattccgga gaatttccta ttctctagaa agtataggaa cttcgacgtc attt 54  
 28 <210> SEQ ID NO: 2  
 29 <211> LENGTH: 27  
 30 <212> TYPE: DNA  
 C--> 31 <213> ORGANISM: Artificial  
 33 <220> FEATURE:  
 34 <223> OTHER INFORMATION: Polylinker  
 36 <400> SEQUENCE: 2  
 37 aaattgaatt cgagctcggt acccggg 27  
 40 <210> SEQ ID NO: 3  
 41 <211> LENGTH: 27  
 42 <212> TYPE: DNA  
 C--> 43 <213> ORGANISM: Artificial  
 45 <220> FEATURE:  
 46 <223> OTHER INFORMATION: Sense Strand of Adaptor  
 48 <400> SEQUENCE: 3  
 49 agcttctgca gcagaccgtg catcatg 27  
 52 <210> SEQ ID NO: 4  
 53 <211> LENGTH: 19  
 54 <212> TYPE: DNA  
 C--> 55 <213> ORGANISM: Artificial  
 57 <220> FEATURE:  
 58 <223> OTHER INFORMATION: Antisense Strand of Adaptor  
 60 <400> SEQUENCE: 4  
 61 atgcacgggc tgctgcaga 19  
 64 <210> SEQ ID NO: 5  
 65 <211> LENGTH: 1285  
 66 <212> TYPE: DNA  
 67 <213> ORGANISM: Homo sapiens  
 69 <400> SEQUENCE: 5  
 70 cagaccgtgc atcatgagcc agtttggcat cctgtgcaag acaccaccta aggtgctggt 60

## RAW SEQUENCE LISTING

DATE: 04/24/2001

PATENT APPLICATION: US/09/807,223

TIME: 13:26:16

Input Set : A:\Q63988 Sequence Listing.txt

Output Set: N:\CRF3\04242001\I807223.raw

```

72 gcgccagttc gtggagaggt ttgagaggcc ctctggagag aagattgcct cctgtgcagc 120
74 tgagctgacc tacctgtgct ggatgatacc ccacaacggc acagccatca agagggccac 180
76 ctttatgagc tacaacacca tcattagcaa ctccctgagc ttcgacattg tgaacaagtc 240
78 cctccagttt aaatacaaga ccagagaaggc cacaatcctg gaggcctccc tgaagaaatt 300
80 gattcctgct tgggagttca ccatcatccc ctacaatggc cagaagcacc agtctgatat 360
82 cactgatatt gtgagcagtc tgcaactcca gttcgagtc tctgaggaag ctgacaaggg 420
84 caacagccac agcaagaaga tgctgaaggc cctgctcagt gagggagaaa gcatctggga 480
86 gatcactgag aagatcctga actcctttga gtacacttcc agattcacca agaccaagac 540
88 cttgtaccag ttctgtttcc tggccacctt catcaactgt ggcaggttca gcgacatcaa 600
90 gaatgtggat cccaaatcct ttaaactggt ccagaacaag tacctgggag tgatcatcca 660
92 gtgcctggtg acagagacca agacctctgt gagcaggcac atctacttct tctctgccag 720
94 gggcaggatt gatccactgg tgtacctgga tgagttcctg aggaactctg agccagtgtc 780
96 gaagcgggtg aacaggaccg gcaactcttc cagcaacaag caggagtacc agctgtctaa 840
98 ggacaacctg gtgaggtcct acaacaaagc tttgaagaaa aatgccccct acccaatctt 900
100 tgccatcaag aatggcccta agtcccacat tggcagacac ctgatgacct ccttcctgtc 960
102 catgaagggc ctgacagagc tgaccaatgt tgtgggcaac tggagcgata agcgggcctc 1020
104 tgccgtggcc agaacaacct atactcacca gatcacagca atccctgatc actacttcgc 1080
106 actggtgtct cgttactatg catatgatcc catctccaag gagatgattg cattgaagga 1140
108 tgagaccaac ccaattgagg agtggcagca cattgagcag ctgaagggtg gtgccgaggg 1200
110 cagcattcgc taccctgcct ggaatgggat catttcccag gaggtgctgg actacctgtc 1260
112 ttctacatc aacagacgca tctga 1285

```

## VERIFICATION SUMMARY

DATE: 04/24/2001

PATENT APPLICATION: US/09/807,223

TIME: 13:26:17

Input Set : A:\Q63988 Sequence Listing.txt

Output Set: N:\CRF3\04242001\I807223.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No  
L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:19 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:1  
L:31 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:2  
L:43 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:3  
L:55 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:4